Automated Segmentation of Scleroderma in High Resolution CT Imagery

Richard L. Fortson, Los Alamos National Lab
Dr. David Lynch, National Jewish Center
Dr. John Newell, National Jewish Center

ABSTRACT

In this paper, we describe a system implemented to automatically classify and quantitatively measure the extent of a lung disease called Scleroderma using High Resolution Computed Tomography (HRCT) imagery. Scleroderma is a disease characterized by a slowly developing fibrosis in the lungs of its victims. Early diagnosis of the extent of the disease using CT imagery can be especially difficult, as many of its visual features are weak and very subtle. These visual subtleties can lead to differences in analysis between radiologists when gauging the exact extent of the disease. Not having an absolute standard with which to judge the extent of the disease can hinder evaluations on the effectiveness of new treatments applied during the disease's early stages. We have attempted to remove the subjective component by designing a robust system which aids radiologists in measuring the extent of the disease during it's earliest stages. This system employs a bank of 17 Maximum Likelihood classifiers trained on the variety of tissue types typically seen within Scleroderma HRCT imagery. The system also employs several heuristic constraints. These constraints are used to mimic some of the decision making processes that radiologists typically employ during their analysis. Results of this classifier system are demonstrated on a series of HRCT exams of patients in the early stages of the disease. These results were found to compare favorably with physiological tests performed on these patients. This research was done as a collaborative effort between Los Alamos National Laboratory (LANL) and the Radiology Department at National Jewish Center for Immunology and Respiratory Medicine (NJCIRM).

INTRODUCTION

Scleroderma is a disease which affects many body systems, but is primarily characterized by thickening and tightening of the skin. More women are affected than men, and the dominant age group is between 20 to 50. The lungs are affected in 70-90% of cases, and develop either fibrosis or changes in the blood vessels which lead to increased pressure in the pulmonary arteries. The fibrosis usually starts with an increase in lung fiber density near the posterior (back) regions of the lungs. HRCT imagery frequently shows the initial effects as an amorphous increase in lung opacity with a hazy appearance like "ground glass". Later stages of fibrosis are characterized by the emergence of a network of coarse lines. These coarse lines eventually develop into regions containing large numbers of small cysts. This end-stage effect is sometimes referred to as "honeycombing" and is non-reversible.

The goal of our research is to use High Resolution Computer Tomography (HRCT) imagery to automatically detect and quantify Scleroderma in its early stages. Scleroderma treatments are most effective when applied during these early stages. Automated detection methods can aid doctors in gauging the effectiveness of new treatments over time. At present, the status of the disease is assessed by (1) physiology tests and (2) visual inspection of chest radiographs. However, the visual segmentation of Scleroderma can sometimes be a subjective affair. Automated classification routines were designed to remove human subjectivity from the visual segmentation process. Removing this subjectivity may allow physicians to more accurately monitor the effectiveness of new Scleroderma treatment over time.

The automated classification system we have constructed detects Scleroderma abnormalities through the use of a supervised training scheme. In this method, a classifier is taught to recognize Scleroderma by feeding it example imagery of tissue abnormalities. The classifier learns from these samples by computing and recognizing specific features contained within the training imagery. The features used to train the system are the standard deviation, skewness and kurtosis of the image intensities within local neighborhoods. These three statistical features are computed within small circular neighborhoods around each pixel in the training imagery. The features are then passed to an advanced clustering scheme which creates a Scleroderma detection "codebook". The "codebook" consists of a defined set of clusters within the feature space. These clusters define the locations within the feature space in which the features of the training data seem to "clump about". Once defined, a "codebook" can be applied to imagery which was not part of the original training set for the purposes of automatic classification. Our clustering method is based on a Gaussian Maximum Likelihood model in which clusters are parameterized according to their means and covariances. A bank of 17 tissue-specific classifiers are employed to cover the large variation of tissue types associated with Scleroderma. The regions segmented by these 17 classifiers are combined together to create the final segmented lung image. In addition to using classifiers, the system also uses several heuristic techniques to ensure that only valid Scleroderma regions are identified. A simplified diagram of the system is shown in Figure 1.

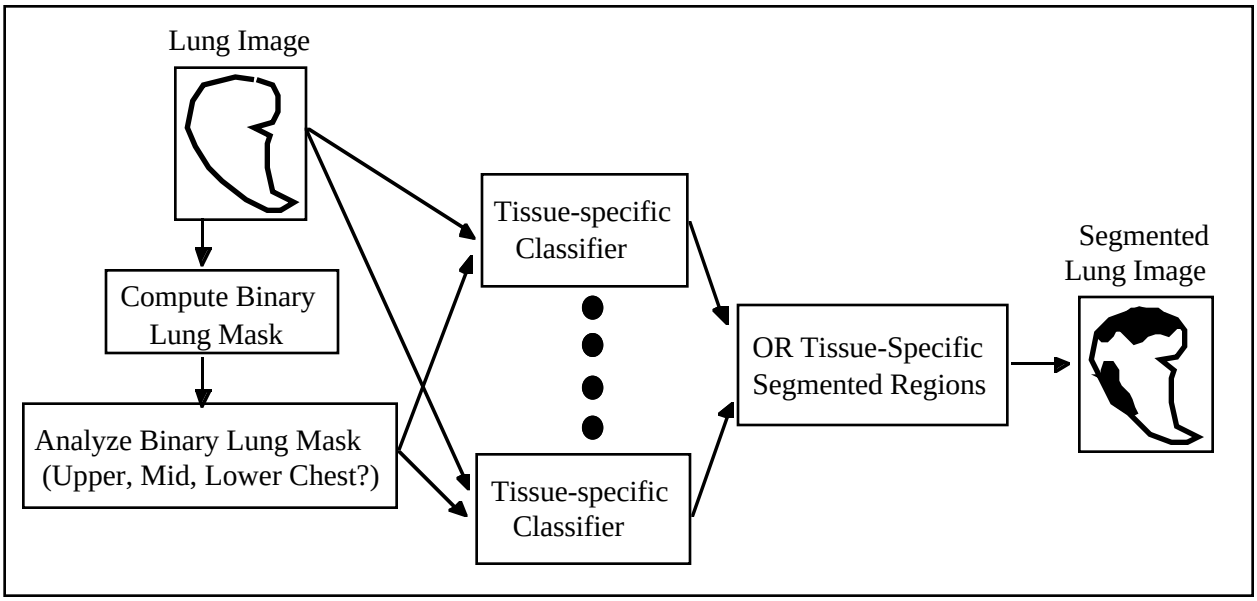


Figure 1. Simplified Auto-Segmentation System Diagram

Our system is designed to segment Scleroderma in re-targeted HRCT lung scenes, where only one lung (left or right) is in prominent view. Re-targeted lung scenes are used because they contain slightly higher (x,y) spatial resolution than full-chest scenes. This higher resolution proves helpful in identifying some of the more subtle visual features of Scleroderma.

PROBLEMS IN SCLERODERMA SEGMENTATION

There are two basic factors which make the classification of Scleroderma a difficult task:

(1) The affected areas have a large *variety* of differences in their appearance in HRCT imagery
(2) The *location* of suspected Scleroderma areas must be taken into account

Both of these factors combine to make measurements of Scleroderma in HRCT imagery a somewhat subjective affair for the radiologist. These two factors must also be addressed within any system which attempts to successfully quantify the extent of the disease.

Scleroderma is a disease which has a large amount of variety in its appearance [1]. This variety in appearance is a function of both differences in the response of a patient's lungs to the disease and upon the particular stage of the disease. Figures 2 through 5 demonstrate this variety by showing four different patients at four different stages of the disease (mild to very severe). In it's earliest stages (Figure 2), the disease is characterized by the emergence of a weak reticulated network of abnormal lines and haze, usually found along the back of the patient's lungs. As the disease progresses, the abnormal lines and haze become stronger and begin to spread deeper into the interior of the lungs (Figure 3). Finally, the system of abnormal lines begin to interconnect and form a pattern of small, empty holes, sometimes referred to as "honeycombs" (Figure 4). At this stage of the disease, the damage is irreversible. As further "honeycombing" continues, the size of the holes can increase and spread to the point that almost the whole lung has been enveloped (Figure 5).

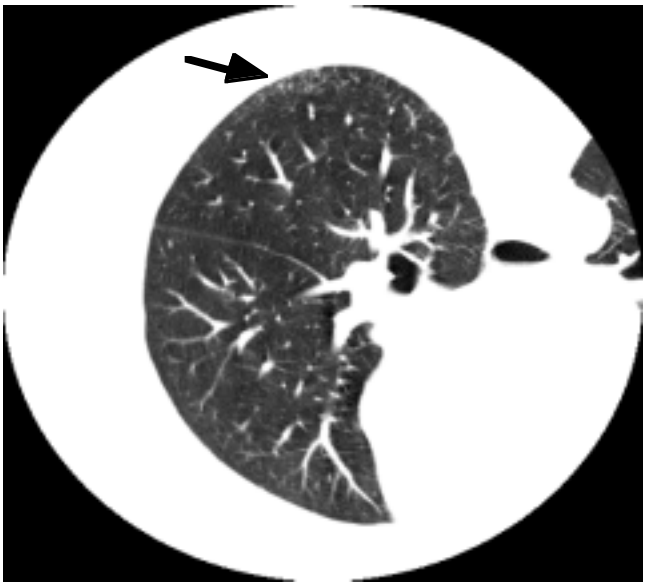


Figure 2. Weak Haze and Abnormal Lines

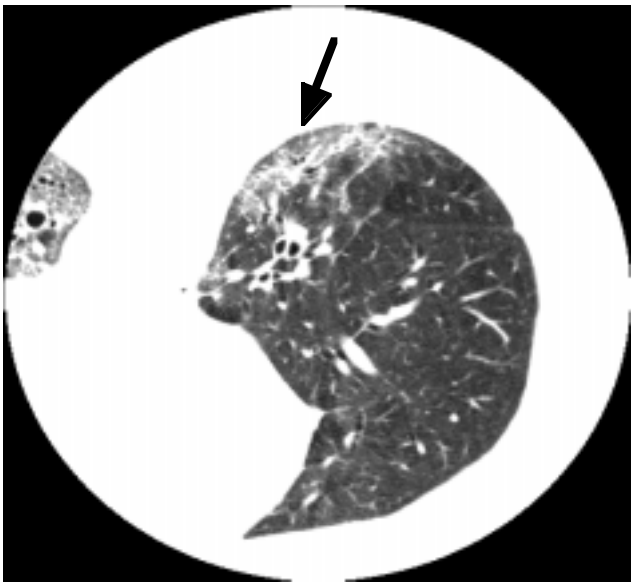


Figure 3. Stronger Haze and Abnormal Lines

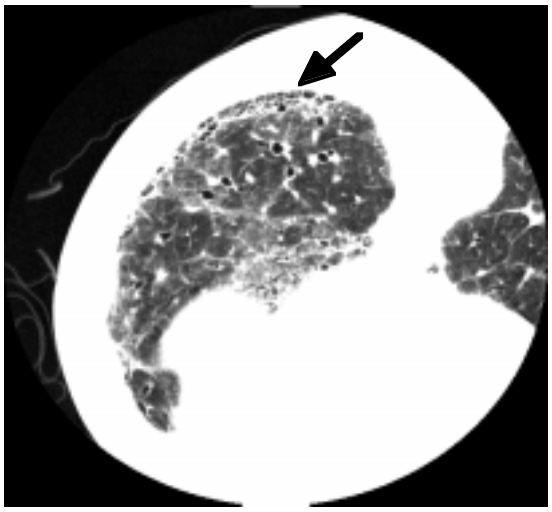


Figure 4. Emergence of Honeycombs

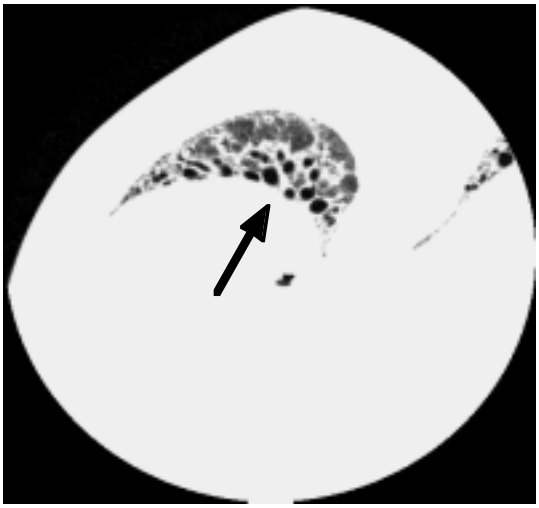
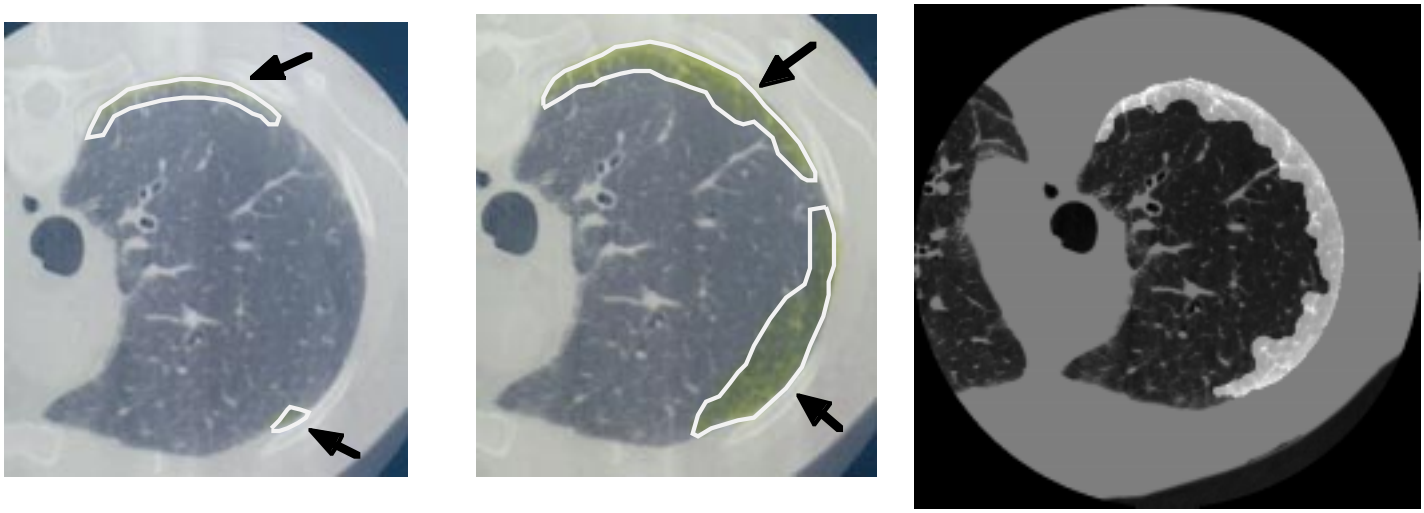


Figure 5. Large Honeycombs

Treatment is most effective if it is applied during the very earliest stages of the disease. However, measuring the extent of Scleroderma in the early stages is a subjective affair, even for trained radiologists. Figure 6 shows a lung in the early stages of the disease where areas of Scleroderma have been delineated by two radiologists. Notice that, while there is some rough agreement on the areas in which Scleroderma exists, there is a large amount of disagreement regarding the total extent of the disease. Results of a computer-based segmentation are shown in this figure. While none of these segmentation decisions can clearly be classified as "wrong", the obvious differences in the human-based segmentations serve to point out the large element of subjectivity inherent in this task. Furthermore, any improvements in the lung through the use of new treatments may initially make only minor changes in the lung's visual appearance. Therefore, having a consistent, non-subjective system for performing the visual analysis might result in a better means for monitoring both the severity of the disease and the effectiveness of any new treatments.



Radiologist 1 **Radiologist 2** **Auto-Segmentation**

Figure 6. Differences in Analysis Between Two Radiologist and Computer Auto-Segmentation

Large variety in the appearance of the affected tissue and the weak and subtle features which are present in the early stages of the disease are only some of the difficulties associated with it's detection. The location of suspected anomalies must also be taken into account. For example, one early sign of Scleroderma's presence is the existence of a reticulated pattern of weak lines and haze, due to stiffening of the lung fibers (fibrosis). However, it is also known that fake abnormal lines and haze can appear in healthy lungs simply due to the compression effects of gravity. When a patient is scanned in the prone (face down) position, gravity causes compression on the forward surfaces of the lungs. This compression is due to the weight of the supporting material which lies above the forward sections of the lung. This compression can cause even normal lung fibers to visually appear as if they are in a state of fibrosis. Therefore, any abnormal lines and haze which appear to exist in the forward lung areas of a prone HRCT scan must be carefully scrutinized before being identified as fibrosis. Another area of caution exists within mid-chest lung surfaces that are adjacent to the heart. Here, the pumping motion of the heart can induce a false haze on the lung surfaces closest to the heart region. Haze can also be induced by the high rates of inter-slice curvature along lung surfaces in the regions near the diaphragm. In this case, the haze is being produced by a volumetric imaging phenomena called "partial-pixel mixing". High rates of local curvature along the scan direction are causing some of the pixels in a collimated slice to consist of a mixture of both lung and non-lung material. Hazy areas are also sometimes seen adjacent to the large white blood vessels present in mid-chest lung imagery. In all of these regions, the radiologist will frequently apply a stricter criteria before classifying these areas as Scleroderma. Examples of these artifacts are depicted in Figure 7.

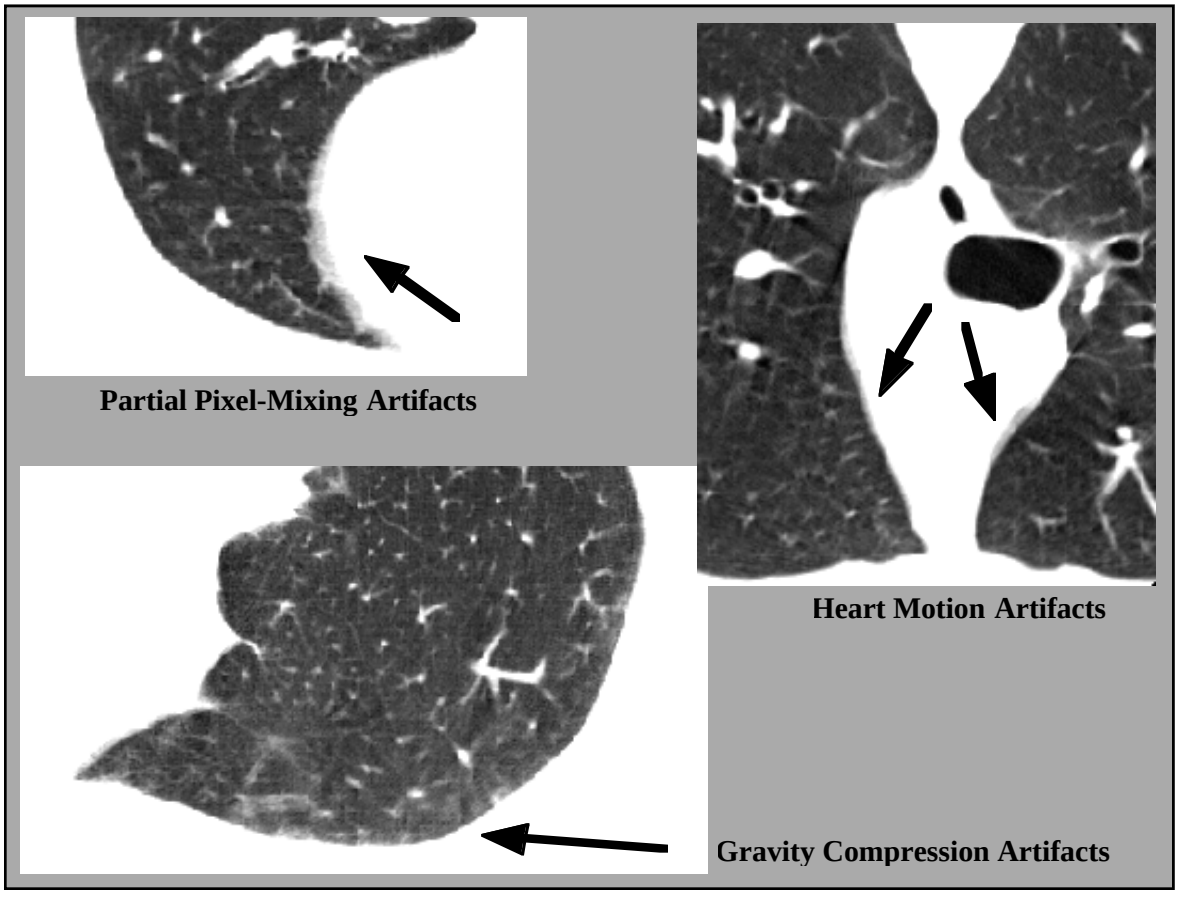


Figure 7. Examples of Artifacts in HRCT Lung Imagery

The previously mentioned regions are those in which the radiologist might show a *negative* location bias. However, there are also regions in which the radiologist might, conversely, show a *positive* location bias. These regions are known Scleroderma "breeding grounds". They can be found along the back of the lungs and along lung surfaces which are adjacent to the left and right sides of the chest wall . Here, regions which may only marginally satisfy the Scleroderma tissue criteria are much more likely to be labeled as Scleroderma by the radiologist.

HOW WE OVERCOME THESE PROBLEMS

Variety in Scleroderma's Appearance

The first problem discussed in the previous section dealt with the large variety in the visual appearance of Scleroderma in HRCT imagery. This problem can be successfully addressed by designing a system which contains a large number of carefully chosen tissue-specific classifiers. Our current system uses a bank of 17 different classifiers, where each classifier belongs to one of the three basic tissue categories inherent to Scleroderma (haze, abnormal lines and honeycombs). Some of the training samples used for these tissue classifiers are shown in Figure 8.

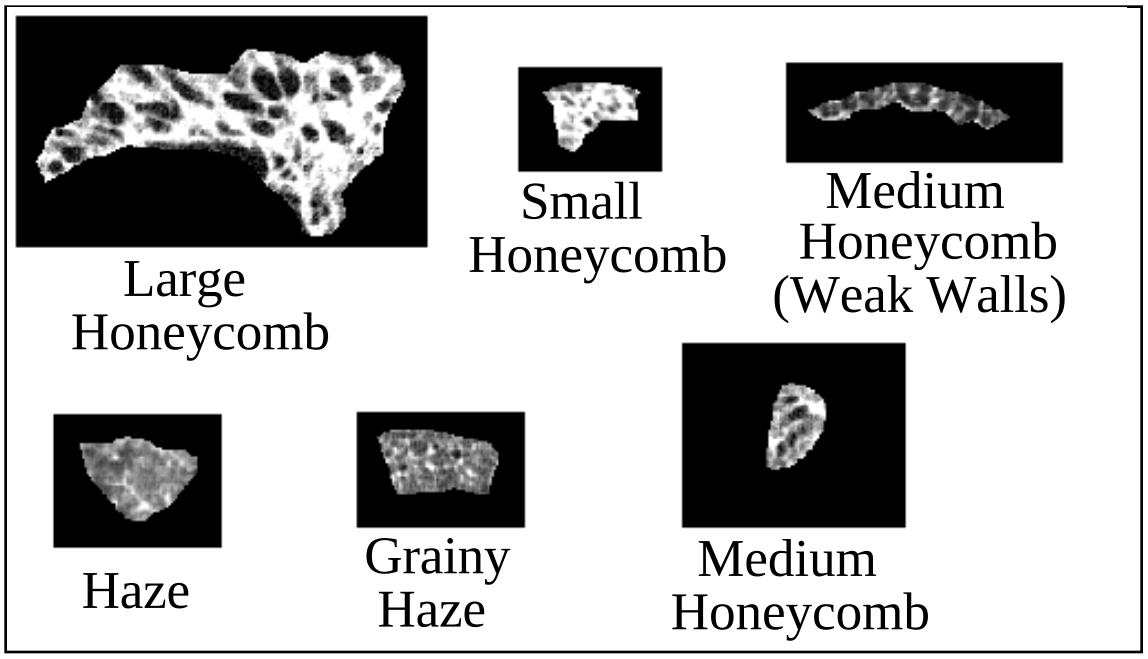


Figure 8. Some Tissue Training Samples Used by the Classifiers

Location Bias

Incorporating a sense of location bias into the system is handled using several different techniques. The most obvious location bias involves identifying and processing only regions which exist within the lung. This is done by creating a binary lung mask. Classifier processing

can only take place within this binary lung mask. Other location bias considerations can be incorporated by:

(1) Removing selected problem regions from the binary lung mask.

This removes certain problem areas from consideration, such as possible hazy regions adjacent to large blood vessels in mid-chest lung imagery.

(2) Creating regional weight masks to apply to the classifier similarity measurements.

This weights the Maximum Likelihood classifier outputs, so that problematic locations will need to meet higher threshold criteria before they will be labeled as "abnormal".

(3) Using a stricter size criteria when deleting very small segmented blobs.

This requires segmented blobs from problematic regions to have a larger size than blobs from other regions before they will be allowed to pass through the system.

In the discussions which follow, we will assume only prone (patient face-down) HRCT scans, so that the top of the image refers to the back regions of the lung and the bottom of the image refers to the front regions of the lung.

Creating the Binary Lung Mask

A binary lung mask for each re-targeted HRCT image is automatically generated by a program which begins by searches for the most significant gray region within the center of the image. This gray region is then thresholded at a level of approximately 900 pixel units to create the left or right binary lung mask. All subsequent classifier-based processing takes place only within this binary lung mask. In cases where the left and right lungs are touching, the program automatically identifies the situation as such and splits the two lungs apart so that each lung can be processed separately.

Removing Large Blood Vessel Regions from Mid-Chest Slices

Areas next to the large, white blood vessels prevalent within mid-chest lung scenes are of special interest. The areas adjacent to these blood vessels sometimes contain hazy regions. These hazy regions are normal, but might be incorrectly classified as abnormal if the segmentation system fails to incorporate a sense of location bias. To adjust for this situation, all incoming lung imagery is tagged for special processing when it comes from the mid-chest sections. In these cases, the large blood vessels and their nearby areas are removed from the binary lung mask.

Identification of mid-chest lung imagery is done automatically through shape analysis of the incoming binary lung mask. Each incoming image is a standard 512-by-512 pixels in size. Mid-chest lungs typically have at least a moderate length across the central row and then begin to taper off near the bottom rows. The decision rule which states this can be written:

IF ($\mathbf{L_{w256}} > 100$) AND ($0.90 > \mathbf{L_{w384}} / \mathbf{L_{w256}} > 0.40$) THEN *mid_chest_lung*

where $\mathbf{L_{w256}}$ is the lung width along row 256
$\mathbf{L_{w384}}$ is the lung width along row 384

If an incoming lung image has been identified as coming from the mid-chest region, its binary lung mask is subject to modifications to ensure that all areas adjacent to large blood vessel are removed. These modifications are done using binary morphology operators. A morphological closing operation is first performed on the binary lung mask to fill in all the small black holes and concave regions caused by the large blood vessels. The original binary lung mask is then XORed with the "closed" binary lung mask, so that only large vessel regions are left. These large vessel regions are then dilated and subtracted from the original binary lung mask to create the modified binary lung mask. This modified binary lung mask excludes both the large vessels and their neighborhood from the classification process (Figure 9).

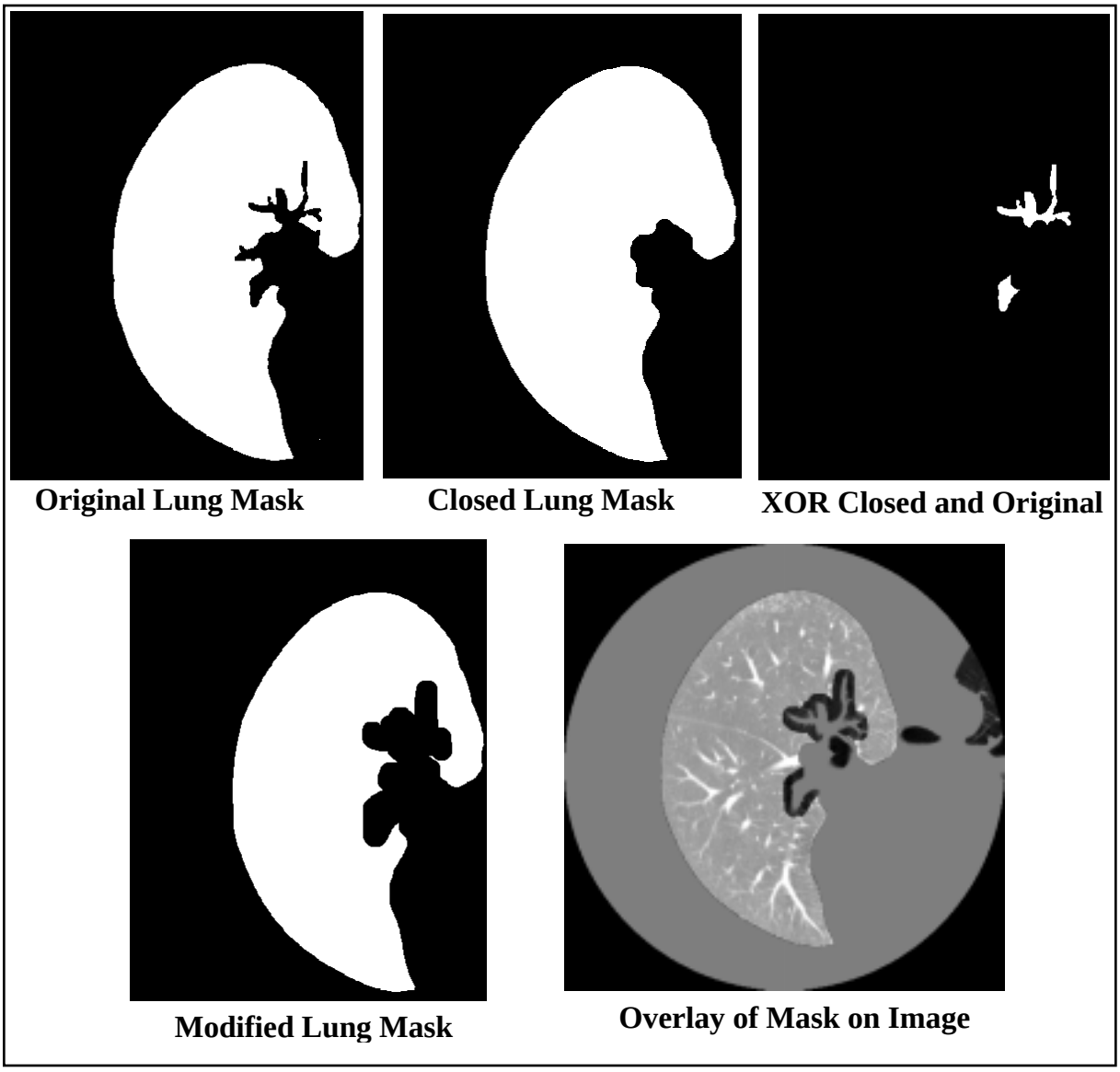


Figure 9. Excluding Large Blood Vessels from the Lung Mask in Mid-Chest Imagery

Compensating for Lung Compression Artifacts Due to Gravity

Gravity compression artifacts are handled by multiplying the similarity measurements output by the Maximum Likelihood classifiers with a gravity gradient mask. The gravity mask is a linear ramp which increases from image top-to-bottom with values ranging from 1.0 to *N* (Figure 10). The exact figure for *N* varies which each classifier, and is a value between 2.0 to 15.0. In general, classifiers which deal with haze and abnormal lines have a much higher *N* value than those which deal with "honeycombing", as the main artifact of gravity compression is to mimic both haze and abnormal lines. It should be noted that the gravity gradient mask does not completely block the segmentation of areas near the bottom of the lung imagery. It simple ensures that tissue samples from these areas will not be classified as Scleroderma unless they are particularly close to the cluster centers defined within each classifier "codebook".

Heart Motion and Pixel-Mixing Artifacts in Mid-Chest Slices

Images identified as coming from mid-chest regions are treated to special processing to compensate for heart motion and "partial pixel-mixing" artifacts. These artifacts exist along certain regions of the lung's periphery and cause a false haze which might be erroneously classified as Scleroderma. To compensate, a special classifier weighting mask is created to suppress classification within these particular areas. This weighting mask consists of floating point values which are added to the cluster similarity measurements output by each Maximum Likelihood classifier (Figure 10). This weighting mask is applied after the application of the previously mentioned gravity gradient mask. The weight values exist only along the frontal and side boundaries of mid-chest lungs. These weight values increase as a function of the row number, so greatest attenuation is near the front surfaces of the lung (corresponding to the bottom regions of the lung image). Attenuation begins at row 400 along the lung boundary facing the chest wall and at row 200 along the lung boundary facing the heart. To determine which lung boundary faces the chest wall and which boundary faces the heart, a determination must be made whether the lung being viewed is from the patient's left or right side. This is done by testing the lung area contained in the bottom two quadrants of the original binary lung mask to determine in which direction the lung appears to be tapering off. For a prone mid-chest HRCT scan, right lungs will tend to taper off towards the bottom-left side of the image and left lungs will tend to taper off towards the bottom-right side. This tapering usually result in a significant difference between the lung areas contained in the two bottom quadrants of the HRCT image. The decision rule states:

IF (0.9 > $\mathbf{L_{a3}} / \mathbf{L_{a4}}$ > 1.1) THEN

  IF ($\mathbf{L_{a3}}$ > $\mathbf{L_{a4}}$) THEN *left_lung*
                ELSE *right_lung*
  ENDIF

ELSE
  undetermined_lung (non-tapered mid-chest lung slice; no attenuation applied)

ENDIF

where $\mathbf{L_{a3}}$ is the lung area in the bottom-left quadrant
$\mathbf{L_{a4}}$ is the lung area in the bottom-right quadrant

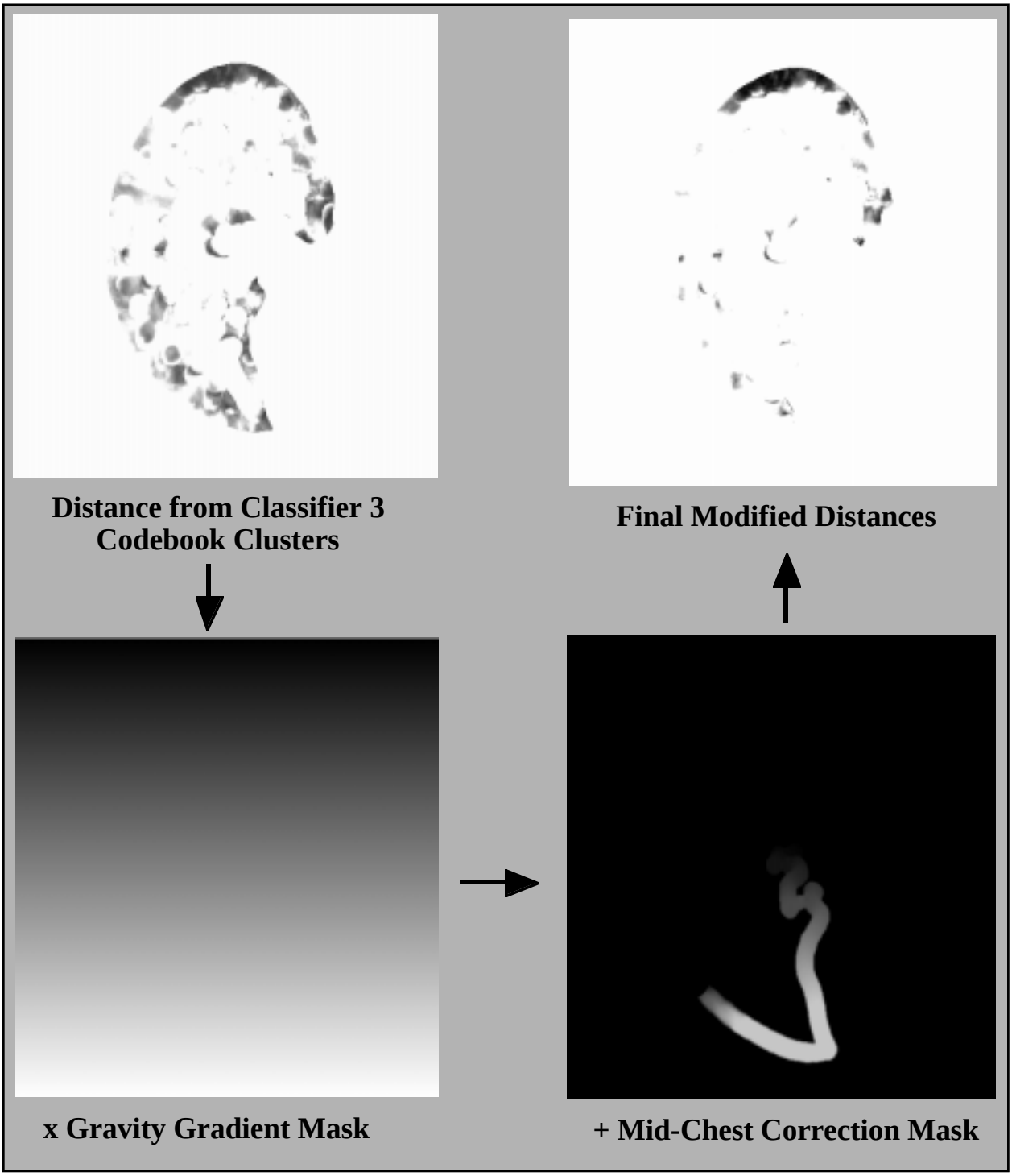


Figure 10. Compensating for Gravity and Heart/Pixel-Mixing Artifacts

Compensating for Favored Scleroderma "Breeding Grounds"

As previously mentioned, Scleroderma tends to have favored "breeding grounds" next to the back of the lungs and along most of the lung's periphery. Segmentation bias towards the back area of the lung is handled using the previously mentioned gravity gradient mask. For prone HRCT scans, this gravity mask causes the system to favor the segmentation of areas towards the back of the chest, where lung compression due to gravity is at a minimum. Bias for regions along the lung periphery is implemented after the cluster distances have been treated to a thresholding operation. This thresholding operation results in the creation of binary blobs (probable Scleroderma regions). All blobs are subjected to a minimum size criteria, so that small "noise" blobs can be removed. Blobs along the lung's periphery are favored over interior blobs by applying a stricter minimum blob size criteria to the interior blobs. Separation of blobs into "edge" and "interior" classifications is done by ORing the edges of the binary lung mask with the binary blobs. A region growing algorithm is then applied which tags all "edge" blobs with a 1 and all "interior" blobs with numbers greater than 1. Once tagged in this way, it is a simple matter to divide the blobs into "edge" and "interior" camps for separate processing to remove small blobs.

THE CLASSIFIER METHODOLOGY

A bank of 17 Scleroderma tissue classifiers are at the heart of this automated segmentation system. It is their task to compute similarity measurements between the features extracted from both the lung and the training sets. The classifiers are based on a clustering mechanism.

In cluster-based classification, a user attempts to find areas in which the extracted features from training data seem to cluster, or "clump about" (Figure 11). The clustering operation involves dividing a set of feature measurements into different non-overlapping groups. Points in a cluster are "more similar" to one another than to the points of other clusters in the feature space. Once the features of the training data have been partitioned into clusters, new data can be classified by extracting the same features and measuring the distance between the feature vectors of the new data and the clusters previously defined by the training data.

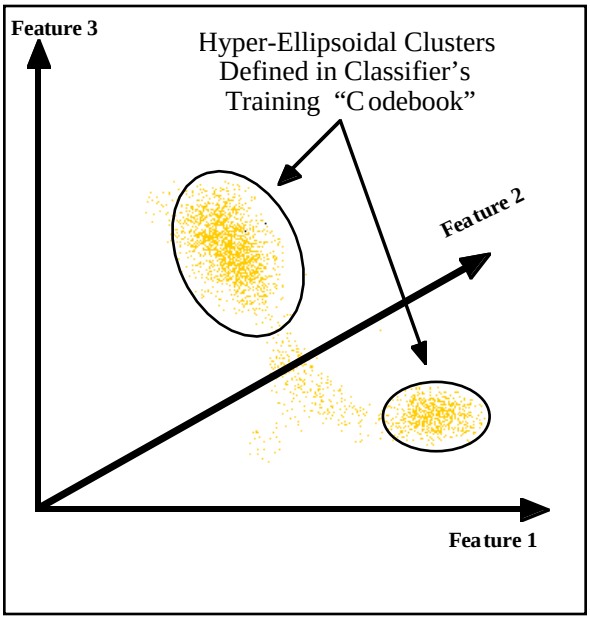


Figure 11. Clustering Features Computed from the Training Imagery

A family of iterative-partitioning algorithms exist for clustering data. These iterative algorithms begin with a set of *k* reference points whose initial values are usually chosen by the user. First, the data points are partitioned into *k* clusters: A data point *x* becomes a member of cluster *i* if z_i, the reference point of cluster *i*, is closest to *x*. The positions of the reference points and the assignment of the data points to clusters are then adjusted during successive iterations. Iterative algorithms are thus similar to fitting routines, which begin with an initial "guess" for each fitted parameter and then optimize their values. Algorithms in this family differ in details of generating and adjusting the partitions. Three members of this family of clustering routines are Lloyd's algorithm, the standard k-means algorithm, and a continuous k-means algorithm first described in 1967 by J. MacQueen [2] and recently developed for general use at Los Alamos.

Initial clustering of the Scleroderma training data is done using an unsupervised, very fast k-means clustering algorithm. The k-means clustering algorithm [3,4] attempts to minimize a squared error cost function by manipulating a set of *k* clusters. We use a version of the nearest neighbor algorithm proposed in [5], where cluster positions are sorted along one of the axes of the data. This algorithm, like many others, does not continue to work effectively as the problem dimension increases. To combat this, we use the first principal component of the data as the axis on which to do our sorting. This axis gives the best possible separation of the data and, thus, allows a faster rate of convergence to the final solution.

After clustering the training data features with a modified k-means algorithm, we then employ a Gaussian Maximum Likelihood algorithm to define the tissue classification "codebook" [6]. In this algorithm, the set of clusters within each of the 17 tissue classifiers are modeled by a set of hyperellipsoids in the feature space. Each cluster hyperellipsoid is defined by a mean vector, $\underline{\mu}$, a covariance matrix, Σ, and an effective radius, d. The mean vector determines the location of the hyperellipsoid, while the covariance matrix determines its shape and orientation in the feature space. Incoming feature vectors are compared to the training clusters by using a Mahanlanobis distance measure, where the distance from an input vector $\underline{x}$ to a cluster mean $\underline{\mu}$ is defined by:

$$(\underline{x} - \underline{\mu}) \quad \Sigma^{-1} \quad (\underline{x} - \underline{\mu})$$

This weighted distance helps to determine whether an incoming vector of features, $\underline{x}$, lies inside or outside of a given hyperellipsoidal cluster. Thresholds for the cluster decision boundaries are set by selecting an appropriate Chi-squared percentile value for each classifier. The exact percentile value used for each tissue classifier varies, but is generally between 90% to 99%.

THE FEATURE SET -- STATISTICAL MOMENTS

Classification is performed using a set of relevant features gleaned from the pixel data. This classification system uses features based on local statistical moments. The feature extraction step begins by remapping the incoming HRCT integer pixel values from 0 to 700 into a byte image. This particular intensity window was chosen because most of the Scleroderma features of interest map well within this range. Circular areas around each pixel in the byte image are then input into the computation of three feature values -- standard deviation, skewness and kurtosis. These three values are loosely based around the concept of statistical moments, where the moment generating function can be expressed as:

$$m_r = \frac{\sum_{j=1}^{N} (X_j - \overline{X})^r}{N}$$

The standard deviation is associated with the 1st moment, skewness with the 2nd moment and kurtosis with the 3rd moment. Standard deviation is simple the measure of dispersion of a distribution. Skewness is the degree of asymmetry of a distribution and is defined as:

$$\text{skewness} = \frac{m_3}{(\sqrt{m_2})^3}$$

Kurtosis is the degree of peakedness of a distribution and is defined as:

$$\text{kurtosis} = \frac{m_4}{m_2^2}$$

The size of the circular windows used for computing theses statistical features varies with the scale of relevant objects contained within each training sample. For example, a training sample consisting of "large honeycomb" might best be served by using a relatively large collection window, say 29 pixels in diameter. This would allow the window to completely cover the large honeycomb objects contained in the sample. Conversely, a "small honeycomb" training sample would best be served by using a smaller collection window, such as one that is 17 pixels in diameter.

The three local statistical features of standard deviation, skewness, and kurtosis seemed to work very well for Scleroderma classification. Prior to using these features, however, an attempt was made at classification using Law's texture features [7,8]. Unfortunately, computing Law's texture features proved time consuming and produced poor results when used as the classifying features. This is probably due to the fact that many of the Scleroderma tissue types lack enough periodicity and structure to create distinctive texture signatures.

THE KHOROS DEVELOPMENT ENVIRONMENT

All routines for this work were developed under the Khoros development system [9]. Khoros is a development environment which provides us with a common platform in which to design our methodologies. There is a very large community of Khoros users, partly because it is an open system available on many different Unix platforms and is distributed free of charge. Khoros programs are typically design as a set of small C code modules which are incorporated into the Khoros *cantata* visual programming language. In *cantata*, each module exist as an separate icon, referred to as a glyph. Sets of glyphs and their interconnections are interactively laid out in a workspace by the user. The visual nature of this environment lends itself to very rapid prototyping of new algorithms and ideas. Figure 12 shows a screen shot of the Khoros workspace which was used in the development of one of the classifiers. Khoros routines can also

be executed via the Unix command line. This allows a set of Khoros routines to also be run from within any Unix shell script -- a handy feature when batch processing is required.

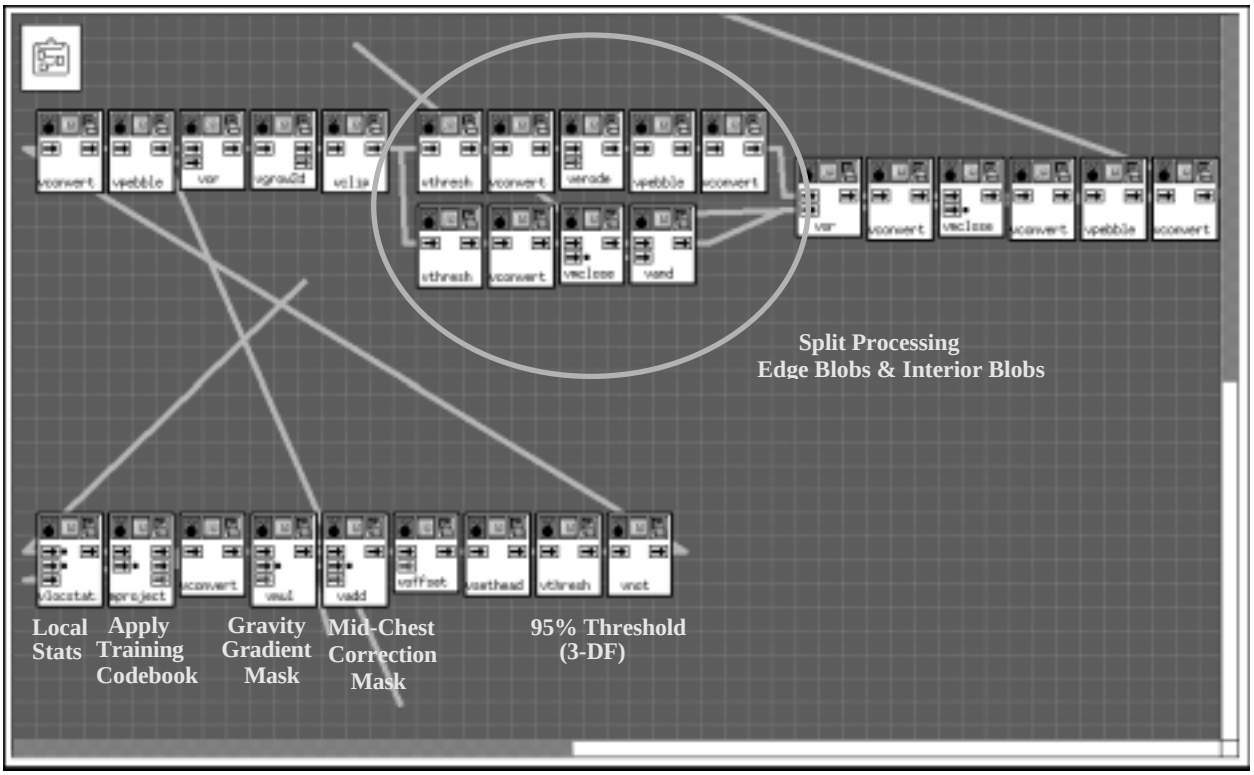


Figure 12. Screen Shoot of Classifier 3 (Grainy Haze) in Khoros Environment

RESULTS OF SCLERODERMA IDENTIFICATION

A set of seven multi-slice HRCT patient exams were supplied for extracting training samples and for testing of the classification system. Each extracted training sample resulted in the construction of a separate classifier "tuned" to detect that particular tissue type. The selection of training samples from among the seven exams were as follows:

Patient 1:	8 Training Samples
Patient 2:	3 Training Samples
Patient 3:	3 Training Samples
Patient 4:	1 Training Sample
Patient 5:	1 Training Sample
Patient 6:	1 Training Sample
Patient 7:	0 Training Samples
---	---
TOTAL :	17 Training Samples

Figures 13 and 14 show some of the results of the Scleroderma segmentation on Patient 7, whose imagery was not contained within any of the training sets. Feedback from two radiologists trained in the detection of Scleroderma has indicated that these segmentation results correspond with the results expected by a trained radiologist.

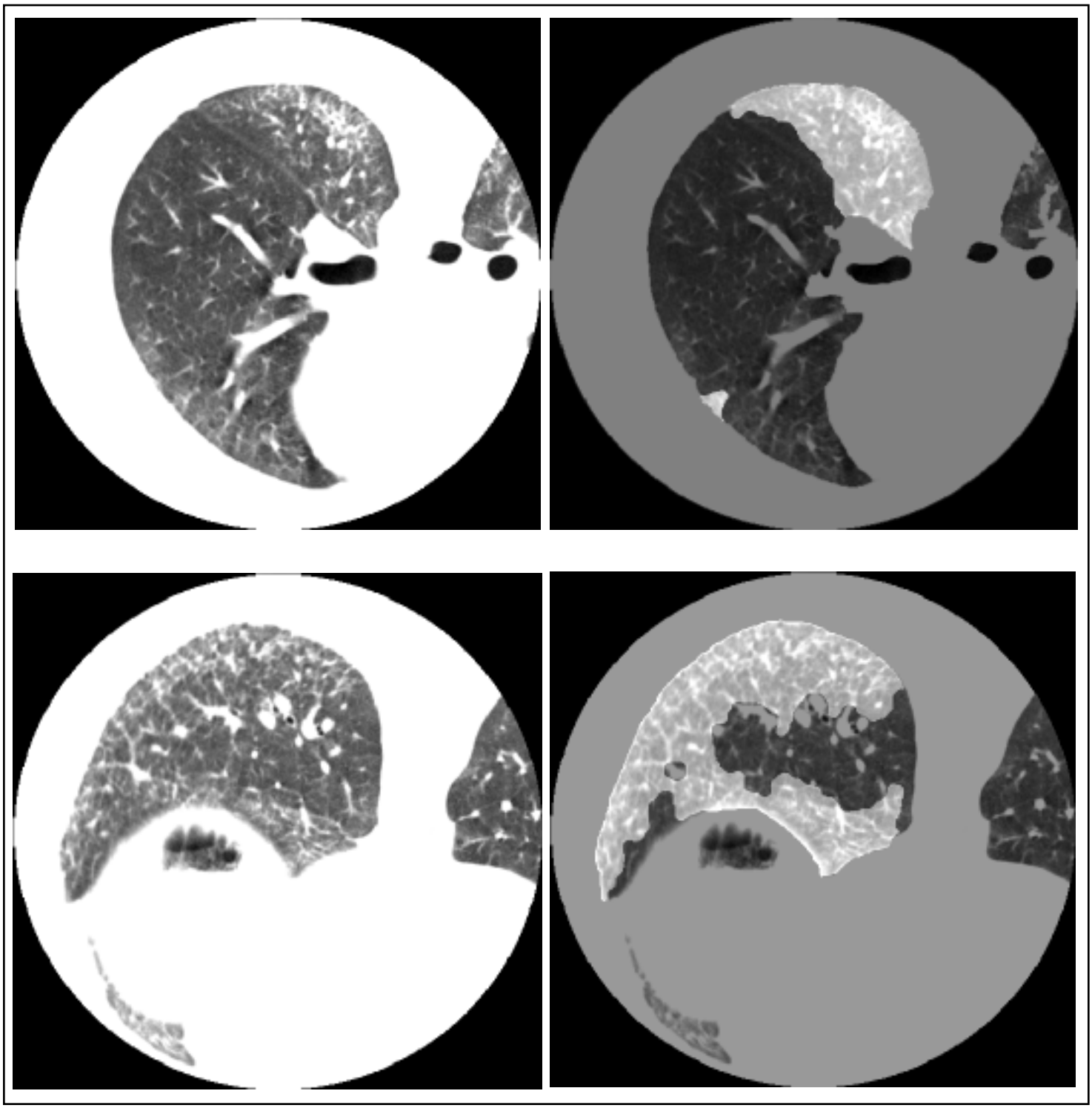


Figure 13. Scleroderma Auto-Segmentation -- Left Lung Slices from Patient 7

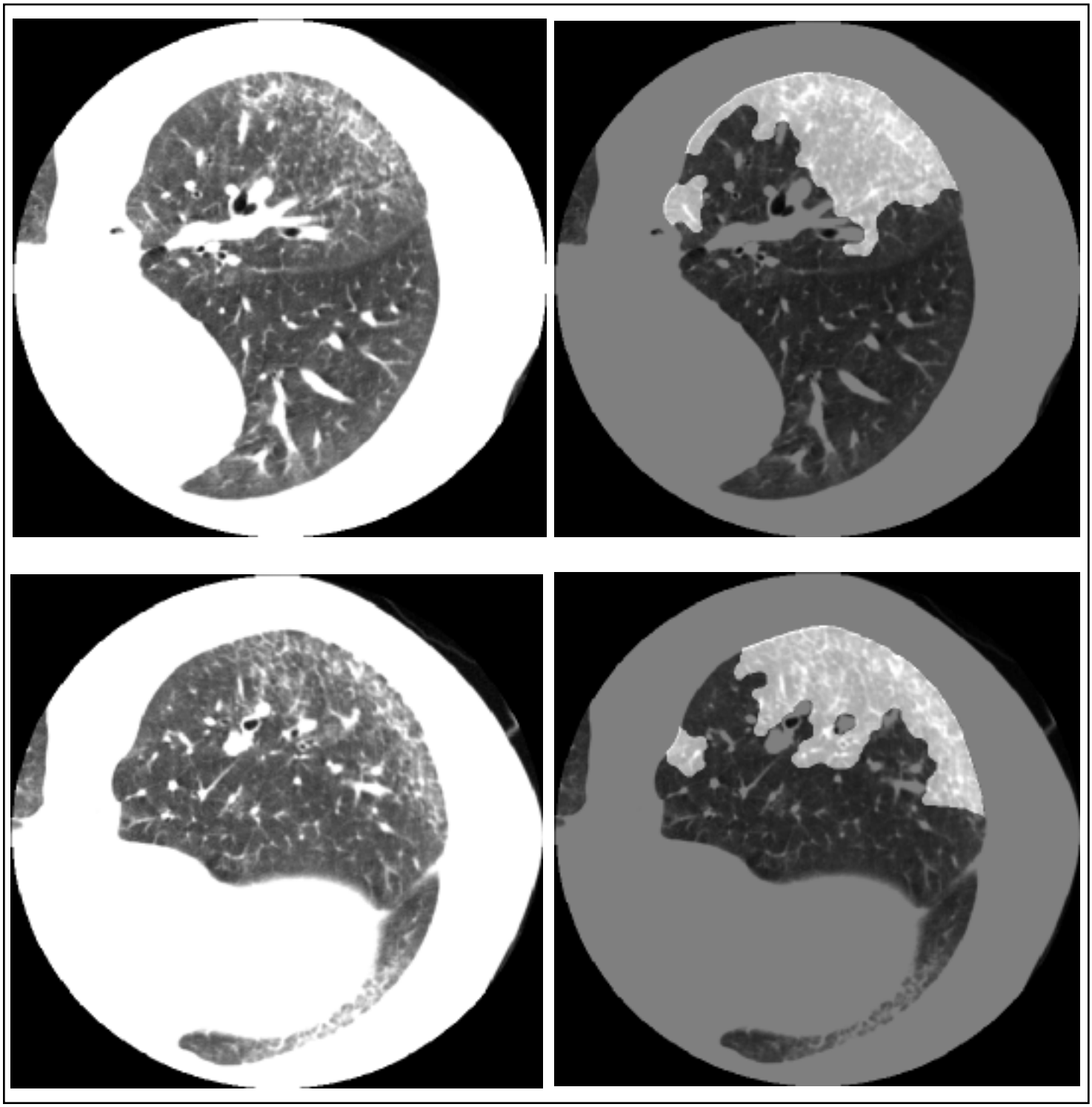


Figure 14. Scleroderma Auto-Segmentation -- Right Lung Slices from Patient 7

To gain further insight into the segmentation results, correlation values were computed between patient physiology data and the total percent of "good" lung detected using the automated segmentation system. Physiology data were available for all patients except Patient 6. The measurements from five physiology tests were utilized.

PREDTLC:
This is a test of the total lung capacity (percent of predicted value). It is a measure of how well the lung can expand, i.e. how stiff is it. It should decrease with increasing fibrotic disease.

PREDFVC:
This is the forced vital capacity (percent of predicted value). Like PREDTLC, it is a measure of lung stiffness.

PBPVCK:
This is the K constant, derived from a pressure-volume curve. It offers a measure of lung compliance which is related to lung stiffness and the degree of fibrosis.

PPDLCO:
This is the lung diffusing capacity (percent of predicted value). It should decrease with increasing fibrosis.

PPDLCOVA:
This is another measure of lung diffusing capacity (percent of predicated value) which has been corrected for the alveolar volume. It should decrease with increasing fibrosis.

Below is a table showing both the physiology data and the "% Good Lung" volume computed using the automated segmentation system for the various patients.

Patient #	PREDTLC	PREDFVC	PBPVCK	PPDLCO	PPDLCOVA	% Good Lung
Patient 1	87.0	62.0	0.08	52.0	73.0	66.22
Patient 2	89.0	61.0	0.28	62.0	84.0	79.13
Patient 3	80.0	67.0	0.10	64.0	100.0	70.52
Patient 4	102.0	91.0	0.14	79.0	83.0	99.41
Patient 5	72.0	50.0	0.11	43.0	80.0	73.34
Patient 7	56.0	40.0	0.14	NA	NA	70.57

The correlation matrix between the five physiology tests and the automated segmentation volumes is as follows:

	% Good Lung	PREDTLC	PREDFVC	PBPVCK	PPDLCO	PPDLCOVA
% Good Lung	1.0000	0.6710	0.8161	0.2704	0.8410	0.1592
PREDTLC	0.6710	1.0000	0.9181	0.1890	0.8535	-0.1513
PREDFVC	0.8161	0.9181	1.0000	0.0035	0.9392	0.1642
PBPVCK	0.2704	0.1890	0.0035	1.0000	0.2557	0.0440
PPDLCO	0.8410	0.8535	0.9392	0.2557	1.0000	0.3738
PPDLCOVA	0.1592	-0.1513	0.1642	0.0440	0.3738	1.0000

As seen from these correlations, the automated segmentation system seems to correlate well with the physiology measurements for PREDTLC, PREDFVC and PPDLCO, with correlation values between 0.67 to 0.84.

FUTURE ENHANCEMENTS

Several enhancements are planned for the present system. One important addition which needs to be considered is the detection of "consolidation masses" within the HRCT imagery. These are regions containing either fluid or very stiff lung fibers. The "consolidation masses" tend to be very bright and very uniform in their appearance. Unfortunately, many normal objects within the HRCT imagery are also very bright and uniform, such as blood vessels. Therefore, features other than the current statistical moments will be necessary. Shape analysis may offer the needed feature for these objects. Another enhancement which we would like to add to the system is the labeling of segmented area as to their broad Scleroderma tissue type (haze, abnormal lines, or honeycomb). This would allow for finer analysis on the exact state of the disease. Finally, it might be helpful to allow the radiologist some gross control over the segmentation mechanism. This could be done by incorporating a five level sensitivity control which ranges from "very insensitive" to "very sensitive". A control such as this would allow the radiologist to create quantitative measurements in a style which suits his or her particular preference, while at the same time allowing consistency between quantitative measurements from different exams based on the same preference setting.

SUMMARY

Advanced clustering methods have proven successful at producing useful quantitative measurements of Scleroderma within HRCT lung imagery. Measurements of the "% Good Lung" output by the system correlate well with many of the available physiology tests. A bank of 17 separate Gaussian Maximum Likelihood classifiers were utilized to capture the large variety of tissue types inherent to this fibrotic disease. Heuristic methods were employed in an attempt to mimic some of the decision-making processes which radiologists typically use when diagnosing a patient.

REFERENCES

[1] W. Webb, N. Muller, D. Naidich. *High Resolution CT of the Lung*. Raven Press, New York, New York 1992.

[2] J. MacQueen. Some methods for classification and analysis of multivariate observations. In *Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability, Volume I, Statistics*. University of California Press.

[3] J.T. Tou and R.C. Gonzales. *Pattern Recognition Principles*. Addison-Wesley, Reading, MA, 1974.

[4] A.K. Jain and R.C. Dubes. *Algorithms for Clustering Data*. Prentice Hall, Englewood Cliffs, NJ, 1988.

[5] J.H. Friedman, F. Baskett, L.J. Shustek. An algorithm for finding nearest neighbors. *IEEE Transactions on Computers*, pages 1000-1006, October 1975.

[6] W. Niblack. *An Introduction to Digital Image Processing*. Prentice Hall, Englewood Cliffs, NJ, 1986.

[7] K. Laws. *Textured Image Segmentation.* Ph.D. dissertation, Univ. of Southern Calif., January 1980.

[8] K. Laws. Rapid texture identification. In *SPIE Vol. 238 Image Processing for Missile Guidance*, pages 376-380, 1980.

[9] Rasure and Williams. An integrated visual language and software development environment. *Journal of Visual Languages and Computing*, 2:217-246, 1991.